

Relative entropy of steering: On its definition and properties

Eneet Kaur*

Mark M. Wilde*[†]

December 22, 2016

Abstract

In [Gallego and Aolita, *Physical Review X* **5**, 041008 (2015)], the authors proposed a definition for the relative entropy of steering and showed that the resulting quantity is a convex steering monotone. Here we advocate a different definition for relative entropy of steering, based on well grounded concerns coming from quantum Shannon theory. We prove that this modified relative entropy of steering is a convex steering monotone. Furthermore, we establish that it is uniformly continuous and faithful, in both cases giving quantitative bounds that should be useful in applications. We also consider a restricted relative entropy of steering which is relevant for the case in which the free operations in the resource theory of steering have a more restricted form (the restricted operations could be more relevant in practical scenarios). The restricted relative entropy of steering is convex, monotone with respect to these restricted operations, uniformly continuous, and faithful.

1 Introduction

Quantum steering corresponds to the scenario in which two parties, typically called Alice and Bob, share a quantum state, and Alice can have an effect on the state of Bob's system if she performs local measurements on hers [EPR35, Sch35, WJD07, CS16]. For certain quantum states, this effect cannot be explained in a classical way, and such states are said to be steerable [WJD07]. Steerable states are necessarily entangled but do not necessarily violate a Bell inequality [CS16].

Quantum steering is relevant as a resource in the context of one-sided device-independent quantum key distribution [BCW⁺12], in which the goal is to distill secret key between Alice, who does not trust the quantum device provided to her, and Bob, who trusts his quantum device. Motivated by this, Ref. [GA15] developed a resource theory of quantum steering, establishing free states in the resource theory as the unsteerable ones and the free operations as one-way local operations and classical communication (1W-LOCC), which preserve the free states. The same authors also defined a steering monotone to be a function that does not increase on average under 1W-LOCC, they proposed a definition for the relative entropy of steering, and they proved that their proposed quantity is a steering monotone.

The relative entropy of steering proposed in [GA15] can be considered in a game-theoretic context with two players and the pay-off function given by the quantum relative entropy. The relative entropy of steering is a function of an assemblage $\{\hat{\rho}_B^{a,x}\}_{a,x}$, defined to be the set of unnormalized

*Hearne Institute for Theoretical Physics, Department of Physics and Astronomy, Louisiana State University, Baton Rouge, Louisiana 70803, USA

[†]Center for Computation and Technology, Louisiana State University, Baton Rouge, Louisiana 70803, USA

states that result on Bob's system B after Alice performs measurement $x \in \mathcal{X}$ and receives outcome $a \in \mathcal{A}$, where \mathcal{X} and \mathcal{A} are finite alphabets. That is, if Alice and Bob share the state ρ_{AB} and Alice performs a positive operator-valued measure (POVM) $\{\Lambda_a^{(x)}\}_a$ on her system, where $\Lambda_a^{(x)} \geq 0$ and $\sum_a \Lambda_a^{(x)} = I_A$, then the resulting assemblage would be $\{\hat{\rho}_B^{a,x} = \text{Tr}_A([\Lambda_a^{(x)} \otimes I_B]\rho_{AB})\}_{a,x}$. The relative entropy of steering proposed in [GA15] quantifies how distinguishable a given assemblage is from one that has a classical description, in terms of the quantum relative entropy [Ume62]. In particular, let us say that Player 1's goal is to maximize the quantum relative entropy between the two assemblages, and he is allowed to perform any 1W-LOCC operation in order to do so. Player 2's goal is to minimize the quantum relative entropy by picking an assemblage that has a classical description. Clearly, we have to pick an order in which the players take their turns. In [GA15], the authors had Player 2 go first, and then Player 1 next. This means that Player 1 can react to the strategy of Player 2, and in particular that the quantity in mathematical terms looks like (crudely)

$$\inf_{\text{LHS}} \sup_{1\text{WLOCC}} D, \quad (1.1)$$

where LHS is the set of assemblages having a classical description, 1WLOCC is the set of 1W-LOCC operations, and D is the quantum relative entropy payoff function (we will define all of this in much more detail later).

The main purpose of the present paper is to advocate for a different definition of the relative entropy of steering in which the order of play described above is exchanged, so that (crudely), the quantity we are proposing is

$$\sup_{1\text{WLOCC}} \inf_{\text{LHS}} D. \quad (1.2)$$

The interpretation is thus that Player 1 first acts to maximize D by “playing” a 1W-LOCC operation, to which Player 2 can react by “playing” an assemblage having a classical description. Our alternate definition for the relative entropy of steering might seem like a minor modification, but we offer three compelling reasons for our proposal:

1. The optimization order for the quantity in (1.2) is consistent with all previously known information-theoretic measures of dynamic resources as considered in quantum Shannon theory [Wil16], including Holevo information of a channel [Hol06], mutual information of a channel [AC97, BSST02], coherent information of a channel [SN96], squashed entanglement of a channel [TGW14], Rains information of a channel [TWW14], etc.
2. The quantity in (1.2) is never larger than that in (1.1) (due to the order of optimizations), and given that the main application of relative entropic quantifiers in quantum Shannon theory has been to get tight upper bounds on distillable entanglement or secret key [Rai01, HHHO05, HHHO09, TWW14, TBR16, PLOB15, WTB16], we suspect that the quantity in (1.2) will be the right one to use in future applications.
3. The game-theoretic interpretation from [vDGG05] would say that (1.2) quantifies the statistical strength of Player 1 to convince Player 2 that the underlying assemblage is steerable, and thus represents a stronger measure or proof of the statistical strength of steerability than does (1.1).

We elaborate more on the first point in Section 3.1.

In the remainder of the paper, we review some preliminaries in Section 2 and provide a formal definition for our proposed relative entropy of steering in Section 3 (we refer to this quantity simply as “the relative entropy of steering” in the remainder of the paper). In Section 3.2, we prove that the relative entropy of steering is a steering monotone, and in Section 3.3 we prove that it is a convex function of the assemblage for which it is evaluated. Thus, the relative entropy of steering is a convex steering monotone according to [GA15, Definition 2]. Section 3.4 establishes upper bounds on the relative entropy of steering. In Section 3.5, we define a metric for assemblages (“trace distance of assemblages”), and we prove that the relative entropy of steering is uniformly continuous with respect to this metric (we give quantitative continuity bounds). In Section 3.6, we prove that the relative entropy of steering is faithful, and we give quantitative faithfulness bounds.

As discussed in [KWW16], we can consider a restricted class of 1W-LOCC operations which might have more relevance in practical scenarios, in which classical communication from Bob to Alice reaches Alice only after she obtains the output of her black box. With this in mind, we define a restricted relative entropy of steering, and we prove that it is a restricted steering monotone, faithful, and uniformly continuous with respect to a metric relevant for restricted 1W-LOCC.

2 Preliminaries

In the introduction, we discussed assemblages as arising from a local measurement of Alice on a bipartite state that she shares with Bob. However, the common approach in the steering literature [CS16], also known as the one-sided device-independent approach, is to consider an assemblage on its own, being defined as a set $\{\hat{\rho}_B^{a,x}\}_{a,x}$ of arbitrary positive semi-definite operators constrained by the no-signaling principle. From the one-sided device-independent perspective, we think of Alice’s system as being a black box, taking a classical input $x \in \mathcal{X}$ and producing a classical output $a \in \mathcal{A}$, where \mathcal{X} and \mathcal{A} are finite alphabets. The no-signaling principle is that the reduced state of Bob’s system should not depend on the input x to Alice’s black box if the output a is not available to him:

$$\sum_a \hat{\rho}_B^{a,x} = \sum_a \hat{\rho}_B^{a,x'} \quad \forall x, x' \in \mathcal{X}. \quad (2.1)$$

We can then define $\rho_B := \sum_a \hat{\rho}_B^{a,x}$ and the last constraint on an assemblage is that ρ_B is a quantum state. With this last constraint, we see that $\text{Tr}(\hat{\rho}_B^{a,x})$ can be interpreted as a conditional probability distribution $p_{\hat{A}|X}$, so that $p_{\hat{A}|X}(a|x) = \text{Tr}(\hat{\rho}_B^{a,x})$.

As discussed in [KWW16], one can think of an assemblage as being similar to a quantum broadcast channel [YHD11], accepting a classical input x from a sender and producing a classical output a with probability $\text{Tr}(\hat{\rho}_B^{a,x})$ for one receiver and a quantum output $\hat{\rho}_B^{a,x} / \text{Tr}(\hat{\rho}_B^{a,x})$ for the other receiver if $\text{Tr}(\hat{\rho}_B^{a,x}) \neq 0$. However, this perspective is not fully complete, given that the quantum system B is accessible to Bob before the input x is chosen. In any case, we say that an assemblage is a *dynamic resource* in the sense of [DHW08], in that its behavior is modified depending on the input x .

An assemblage is *unsteerable* if arises from a classical, shared random variable Λ in the following sense [WJD07]:

$$\hat{\rho}_B^{a,x} = \sum_{\lambda} p_{\Lambda}(\lambda) p_{A|X\Lambda}(a|x, \lambda) \rho_B^{\lambda}, \quad (2.2)$$

where $p_{\Lambda}(\lambda)$ is a probability distribution for Λ , $p_{A|X\Lambda}$ is a conditional probability distribution, and ρ_B^{λ} is a quantum state. The above structure indicates that the correlations observed can be

explained by a classical random variable Λ , a copy of which is sent to both Alice and Bob, who then take actions conditioned on a particular realization λ of Λ . The set of all unsteerable assemblages is referred to as LHS (short for assemblages having a “local-hidden-state model”).

As discussed in the introduction, the most general free operations allowed in the context of quantum steering are 1W-LOCC. As a particular example, starting with a given assemblage $\{\hat{\rho}_B^{a,x}\}_{a,x}$, it is possible for Bob to perform a generalized measurement on his system, specified as the following measurement channel acting on an input state σ_B :

$$\mathcal{M}_{B \rightarrow B'Y}(\sigma_B) := \sum_y K_y \sigma_B K_y^\dagger \otimes |y\rangle\langle y|_Y, \quad (2.3)$$

where $\sum_y K_y^\dagger K_y = I_B$ and each K_y is a Kraus operator, taking a vector in \mathcal{H}_B to a vector in $\mathcal{H}_{B'}$. Also, $\{|y\rangle\}_y$ denotes an orthonormal basis. Bob can then communicate the classical result y to Alice, who chooses the input x to her black box according to a classical channel $p_{X|Y}(x|y)$. The state after these operations is

$$\rho_{X\hat{A}B'Y} := \sum_{a,x,y} p_{X|Y}(x|y) |x\rangle\langle x|_X \otimes |a\rangle\langle a|_{\hat{A}} \otimes K_y \hat{\rho}_B^{a,x} K_y^\dagger \otimes |y\rangle\langle y|_Y, \quad (2.4)$$

where $\{|x\rangle\}_x$ and $\{|a\rangle\}_a$ denote orthonormal bases.

We now recall the definition of quantum relative entropy, one of the main tools used in this paper. The quantum relative entropy $D(\rho\|\sigma)$ accepts two quantum states ρ and σ as input and outputs a non-negative real number. It is defined as [Ume62]

$$D(\rho\|\sigma) := \text{Tr}(\rho [\log_2 \rho - \log_2 \sigma]) \quad (2.5)$$

if the support of ρ is contained in the support of σ and otherwise it is set to $+\infty$. In the above definition, we take the common convention that the operator logarithms are defined on the support of their arguments. The most critical property of quantum relative entropy is that it is monotone with respect to a quantum channel \mathcal{N} [Lin75, Uhl77], in the sense that

$$D(\rho\|\sigma) \geq D(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)). \quad (2.6)$$

The quantum relative entropy obeys the following property:

$$D\left(\sum_x r(x) |x\rangle\langle x| \otimes \lambda^x \left\| \sum_x s(x) |x\rangle\langle x| \otimes \mu^x\right.\right) = \sum_x r(x) D(\lambda^x\|\mu^x) + D(r\|s), \quad (2.7)$$

which holds for probability distributions r and s and sets of density operators $\{\lambda^x\}_x$ and $\{\mu^x\}_x$ and where $\{|x\rangle\}_x$ is an orthonormal basis. Note that if we write $D(p\|q)$ for probability distributions p and q , then it is implicit that these distributions are encoded along the diagonal of a density operator, so that the corresponding states are commuting.

The quantum entropy is defined as $H(G)_\kappa := H(\kappa_G) := -\text{Tr}(\kappa_G \log_2 \kappa_G)$ for a state κ_G on system G .

3 Relative entropy of steering

In this section, we first give our proposed definition of relative entropy of steering. We then show that it is a convex steering monotone. The subsections thereafter establish upper bounds on it, the trace distance of assemblages as a metric on assemblages, uniform continuity of the relative entropy of steering, and its faithfulness.

Definition 1 (Relative entropy of steering) *Let $\{\hat{\rho}_B^{a,x}\}_{a,x}$ denote an assemblage. We define the relative entropy of steering as follows:*

$$R_S(\hat{A}; B)_{\hat{\rho}} := \sup_{\{p_{X|Y}, \{K_y\}_y\}} \inf_{\{\hat{\sigma}_B^{a,x}\}_{a,x} \in \text{LHS}} D(\rho_{X\hat{A}B'Y} \| \sigma_{X\hat{A}B'Y}), \quad (3.1)$$

where

$$\rho_{X\hat{A}B'Y} := \sum_{x,a,y} p_{X|Y}(x|y) |x\rangle\langle x|_X \otimes |a\rangle\langle a|_{\hat{A}} \otimes K_y \hat{\rho}_B^{a,x} K_y^\dagger \otimes |y\rangle\langle y|_Y, \quad (3.2)$$

$$\sigma_{X\hat{A}B'Y} := \sum_{x,a,y} p_{X|Y}(x|y) |x\rangle\langle x|_X \otimes |a\rangle\langle a|_{\hat{A}} \otimes K_y \hat{\sigma}_B^{a,x} K_y^\dagger \otimes |y\rangle\langle y|_Y, \quad (3.3)$$

LHS denotes the set of all assemblages having a local-hidden-state model, and $\{p_{X|Y}, \{K_y\}_y\}$ denotes a 1W-LOCC operation as described in (2.4).

Remark 2 *By using the property of relative entropy recalled in (2.7), the relative entropy of steering defined in [GA15] can be written as*

$$\inf_{\{\hat{\sigma}_B^{a,x}\}_{a,x} \in \text{LHS}} \sup_{\{p_{X|Y}, \{K_y\}_y\}} D(\rho_{X\hat{A}B'Y} \| \sigma_{X\hat{A}B'Y}), \quad (3.4)$$

with the symbols involved defined as above.

3.1 Justification for Definition 1

We gave three reasons in the introduction that advocate for Definition 1 to be the relative entropy of steering over the definition given in [GA15]. We now elaborate on the first reason, which is that the order of optimizations in Definition 1 is consistent with the order of optimizations given in all known information-theoretic measures of a dynamic quantum resource. Since an assemblage is a dynamic resource as discussed in Section 2, we see no strong reason why the order of optimizations in the relative entropy of steering should not be consistent with all of these other measures.

We first briefly recall some definitions. The quantum mutual information and coherent information of a bipartite state ρ_{AB} can be defined, respectively, as

$$I(A; B)_\rho := \inf_{\sigma_B} D(\rho_{AB} \| \rho_A \otimes \sigma_B), \quad (3.5)$$

$$I(A)B)_\rho := \inf_{\sigma_B} D(\rho_{AB} \| I_A \otimes \sigma_B), \quad (3.6)$$

where the optimizations are with respect to a quantum state σ_B (see, e.g., [Wil16, Section 11.8.1]). The conditional mutual information of a tripartite state ρ_{ABE} can be defined as

$$I(A; B|E)_\rho := I(A; BE)_\rho - I(A; E)_\rho. \quad (3.7)$$

A dynamic resource of primary interest in quantum Shannon theory is a quantum channel $\mathcal{N}_{A \rightarrow B}$, which accepts a state on an input quantum system A and physically transforms it to a state on an output quantum system B . One of the main goals of quantum Shannon theory is to determine capacities of a quantum channel for various communication tasks. The result of many years of effort is that different functions of a quantum channel characterize its different capacities. For example, the classical capacity is characterized by the Holevo information [Hol98, SW97, Hol06], the entanglement-assisted capacity by the channel's mutual information [BSST99, BSST02], and the quantum capacity by the channel's coherent information [Llo97, Sho02, Dev05], respectively defined as

$$\sup_{\rho_{XA}} I(X; B)_{\mathcal{N}(\rho)} = \sup_{\rho_{XA}} \inf_{\sigma_B} D(\mathcal{N}_{A \rightarrow B}(\rho_{XA}) \| \rho_X \otimes \sigma_B), \quad (3.8)$$

$$\sup_{\rho_{RA}} I(R; B)_{\mathcal{N}(\rho)} = \sup_{\rho_{RA}} \inf_{\sigma_B} D(\mathcal{N}_{A \rightarrow B}(\rho_{RA}) \| \rho_R \otimes \sigma_B), \quad (3.9)$$

$$\sup_{\rho_{RA}} I(R)_{\mathcal{N}(\rho)} = \sup_{\rho_{RA}} \inf_{\sigma_B} D(\mathcal{N}_{A \rightarrow B}(\rho_{RA}) \| I_R \otimes \sigma_B). \quad (3.10)$$

In the first line, there is a constraint that system X is a classical system while system A is quantum. In the last two expressions, systems R and A are quantum. The expressions on the right-hand side indicate that the information quantities can be thought of as a comparison between the output of the actual channel and the output of a useless channel, which is one that traces out the input system A and replaces it with the state σ_B . We see in each case that the order of optimization is critically taken to be such that the maximizing player goes first, inputting a state intended to give the best possible discrimination between the channel $\mathcal{N}_{A \rightarrow B}$ of interest and a useless channel. The minimizing player goes second, being able to react to the play of the maximizer by choosing the worst possible useless channel depending on the state $\mathcal{N}_{A \rightarrow B}(\rho_{XA})$ or $\mathcal{N}_{A \rightarrow B}(\rho_{RA})$.

Other information measures which have been used to give upper bounds on communication tasks include the squashed entanglement of a channel [TGW14], the Rains information of a channel [TWW14], and a channel's relative entropy of entanglement [TWW14, PLOB15, WTB16]. These are defined respectively as

$$\sup_{\psi_{RA}} \inf_{\mathcal{S}_{E \rightarrow E'}} I(A; B|E')_{\omega}, \quad (3.11)$$

$$\sup_{\rho_{RA}} \inf_{\sigma_{AB} \in \text{PPT}'} D(\mathcal{N}_{A \rightarrow B}(\rho_{RA}) \| \sigma_{AB}), \quad (3.12)$$

$$\sup_{\rho_{RA}} \inf_{\sigma_{AB} \in \text{SEP}} D(\mathcal{N}_{A \rightarrow B}(\rho_{RA}) \| \sigma_{AB}), \quad (3.13)$$

where in the squashed entanglement of a channel, we take $\omega_{ABE'} := \mathcal{S}_{E \rightarrow E'}(\mathcal{U}_{A \rightarrow BE}^{\mathcal{N}}(\psi_{RA}))$, with ψ_{RA} a pure state, $\mathcal{U}_{A \rightarrow BE}^{\mathcal{N}}$ a fixed isometric extension of the channel $\mathcal{N}_{A \rightarrow B}$, and $\mathcal{S}_{E \rightarrow E'}$ a channel known as a squashing channel. In the latter two lines, PPT' is a set of subnormalized states related to and containing the positive-partial-transpose (PPT) states, and SEP denotes the set of separable, unentangled states. Thus, the interpretation is the same as above: an input to the channel is chosen and then an adversary reacts to this input by trying to minimize the discrimination measure. Note that the latter two quantities have found application as upper bounds on quantum capacity and private capacity, in part because they involve a comparison with a PPT state, which is useless for quantum communication [Rai01], and with a separable state, which is useless for private communication [CLL04, HHHO05, HHHO09].

Thus, given the above list of information measures which have found extensive use throughout quantum Shannon theory and given that each of them have the optimization order as $\sup \inf$, we suspect that this optimization order will be the right approach to take for the relative entropy of steering. Note also that, similar to all of the above information measures, the relative entropy of steering involves a comparison between a given assemblage and another which is useless in the context of steering, in the sense that the latter has a local-hidden-state model and is thus unsteerable.

3.2 Steering monotone

We now prove that the relative entropy of steering is a steering monotone, however deferring the faithfulness proof until Section 3.6:

Theorem 3 (Steering monotone) *Let $\{\hat{\rho}_B^{a,x}\}_{a,x}$ be an assemblage, and suppose that*

$$\left\{ \hat{\rho}_{B_f, z}^{a_f, x_f} := \sum_{a, x} p(a_f | x_f, x, a, z) p(x | x_f, z) K_z \hat{\rho}_B^{a, x} K_z^\dagger / p(z) \right\}_{a_f, x_f}, \quad (3.14)$$

is an assemblage that arises from it by the action of a general 1W-LOCC operation [GA15, Definition 1], where

$$p(z) := \text{Tr} \left(K_z^\dagger K_z \sum_a \hat{\rho}_B^{a, x} \right) = \text{Tr}(K_z^\dagger K_z \rho_B). \quad (3.15)$$

Then

$$\sum_z p(z) R_S(\hat{A}_f; B_f)_{\hat{\rho}_z} \leq R_S(\hat{A}; B)_{\hat{\rho}}. \quad (3.16)$$

Proof. Let $\{\hat{\sigma}_B^{a,x}\}_{a,x}$ be an LHS assemblage, and suppose that

$$\left\{ \hat{\sigma}_{B_f, z}^{a_f, x_f} := \sum_{a, x} p(a_f | x_f, x, a, z) p(x | x_f, z) K_z \hat{\sigma}_B^{a, x} K_z^\dagger / q(z) \right\}_{a_f, x_f}, \quad (3.17)$$

is an LHS assemblage that arises from it by the action of the same 1W-LOCC operation as above, where

$$q(z) := \text{Tr} \left(K_z^\dagger K_z \sum_a \hat{\sigma}_B^{a, x} \right) = \text{Tr}(K_z^\dagger K_z \sigma_B). \quad (3.18)$$

The assemblage $\{\hat{\sigma}_{B_f, z}^{a_f, x_f}\}_{a, x}$ is guaranteed to be an LHS assemblage by [GA15, Theorem 1]. Consider that, in accordance with the definition of $R_S(\hat{A}_f; B_f)_{\hat{\rho}_z}$, the assemblages $\{\hat{\rho}_{B_f, z}^{a_f, x_f}\}_{a_f, x_f}$ and $\{\hat{\sigma}_{B_f, z}^{a_f, x_f}\}_{a_f, x_f}$ can be further preprocessed by a z -dependent 1W-LOCC $\{p_{X_f|YZ=z}, \{L_y^{(z)}\}_y\}$, resulting in the following state:

$$\omega_{X_f \hat{A}_f B'_f Y}^z := \sum_{a_f, x_f, y} p(x_f | yz) [x_f] \otimes [a_f] \otimes L_y^{(z)} \hat{\rho}_{B_f, z}^{a_f, x_f} (L_y^{(z)})^\dagger \otimes [y], \quad (3.19)$$

$$\tau_{X_f \hat{A}_f B'_f Y}^z := \sum_{a_f, x_f, y} p(x_f | yz) [x_f] \otimes [a_f] \otimes L_y^{(z)} \hat{\sigma}_{B_f, z}^{a_f, x_f} (L_y^{(z)})^\dagger \otimes [y]. \quad (3.20)$$

Notation 4 In the above and in what follows, we employ a shorthand $[x] \equiv |x\rangle\langle x|_X$ or $[a] \equiv |a\rangle\langle a|_{\hat{A}}$, etc.

The above states can be embedded in the following ones:

$$\omega_{X_f \hat{A}_f B'_f Y Z} := \sum_z \omega_{X_f \hat{A}_f B'_f Y}^z \otimes p(z)[z], \quad (3.21)$$

$$\tau_{X_f \hat{A}_f B'_f Y Z} := \sum_z \tau_{X_f \hat{A}_f B'_f Y}^z \otimes q(z)[z]. \quad (3.22)$$

The states above are extended by the following ones:

$$\begin{aligned} \omega_{X_f X \hat{A}_f \hat{A} B'_f Y}^z &:= \sum_{a_f, a, x, x_f, y} p(x_f | yz)[x_f] \otimes p(x | x_f, z)[x] \otimes p(a_f | x_f, x, a, z)[a_f] \\ &\quad \otimes [a] \otimes \frac{L_y^{(z)} K_z \hat{\rho}_B^{a, x} K_z^\dagger (L_y^{(z)})^\dagger}{p(z)} \otimes [y], \end{aligned} \quad (3.23)$$

$$\begin{aligned} \tau_{X_f X \hat{A}_f \hat{A} B'_f Y}^z &:= \sum_{a_f, a, x, x_f, y} p(x_f | yz)[x_f] \otimes p(x | x_f, z)[x] \otimes p(a_f | x_f, x, a, z)[a_f] \\ &\quad \otimes [a] \otimes \frac{L_y^{(z)} K_z \hat{\sigma}_B^{a, x} K_z^\dagger (L_y^{(z)})^\dagger}{q(z)} \otimes [y], \end{aligned} \quad (3.24)$$

which in turn are elements of the following classical–quantum states:

$$\omega_{X_f X \hat{A}_f \hat{A} B'_f Y Z} := \sum_z \omega_{X_f X \hat{A}_f \hat{A} B'_f Y}^z \otimes p(z)[z], \quad (3.25)$$

$$\tau_{X_f X \hat{A}_f \hat{A} B'_f Y Z} := \sum_z \tau_{X_f X \hat{A}_f \hat{A} B'_f Y}^z \otimes q(z)[z]. \quad (3.26)$$

Consider that

$$\begin{aligned} &\sum_z p(z) \inf_{\hat{\zeta}^z \in \text{LHS}} D(\omega_{X_f \hat{A}_f B'_f Y}^z \| \zeta_{X_f \hat{A}_f B'_f Y}^z) \\ &\leq \sum_z p(z) D(\omega_{X_f \hat{A}_f B'_f Y}^z \| \tau_{X_f \hat{A}_f B'_f Y}^z) \end{aligned} \quad (3.27)$$

$$\leq \sum_z p(z) D(\omega_{X_f \hat{A}_f B'_f Y}^z \| \tau_{X_f \hat{A}_f B'_f Y}^z) + D(p \| q) \quad (3.28)$$

$$= D(\omega_{X_f \hat{A}_f B'_f Y Z} \| \tau_{X_f \hat{A}_f B'_f Y Z}) \quad (3.29)$$

$$\leq D(\omega_{X_f X \hat{A}_f \hat{A} B'_f Y Z} \| \tau_{X_f X \hat{A}_f \hat{A} B'_f Y Z}) \quad (3.30)$$

$$= D(\omega_{X_f X \hat{A} B'_f Y Z} \| \tau_{X_f X \hat{A} B'_f Y Z}). \quad (3.31)$$

In the first line, we take $\hat{\zeta}^z$ to denote a general LHS assemblage $\{\hat{\zeta}_{B_f}^{a_f, x_f, z}\}_{a_f, x_f}$ and $\zeta_{X_f \hat{A}_f B'_f Y}^z$ denotes the following state:

$$\zeta_{X_f \hat{A}_f B'_f Y}^z := \sum_{a_f, x_f, y} p(x_f | yz)[x_f] \otimes [a_f] \otimes L_y^{(z)} \hat{\zeta}_{B_f}^{a_f, x_f, z} (L_y^{(z)})^\dagger \otimes [y]. \quad (3.32)$$

The first inequality follows by considering that the state $\tau_{X_f \hat{A}_f B'_f Y}^z$ arises from the action of the z -dependent 1W-LOCC operation $\{p_{X_f|YZ=z}, \{L_y^{(z)}\}_y\}$ on the LHS assemblage $\{\hat{\sigma}_{B_f, z}^{a_f, x_f}\}_{a_f, x_f}$. The second inequality follows from non-negativity of relative entropy. The first equality is a consequence of the property of relative entropy recalled in (2.7). The final inequality follows from the data processing inequality for quantum relative entropy, and the final equality follows because the random variable in \hat{A}_f for each state $\omega_{X_f X \hat{A}_f \hat{A}_{B'_f} Y Z}$ and $\tau_{X_f X \hat{A}_f \hat{A}_{B'_f} Y Z}$ is produced by the same classical channel $p(a_f|x_f, x, a, z)$, so that we get the inequality $D(\omega_{X_f X \hat{A}_f \hat{A}_{B'_f} Y Z} \|\tau_{X_f X \hat{A}_f \hat{A}_{B'_f} Y Z}) \leq D(\omega_{X_f X \hat{A}_{B'_f} Y Z} \|\tau_{X_f X \hat{A}_{B'_f} Y Z})$ by data processing and the opposite inequality follows by taking a partial trace over system \hat{A}_f .

We have shown that the above chain of inequalities holds for all assemblages $\{\hat{\sigma}_B^{a, x}\}_{a, x} \in \text{LHS}$, and so we can conclude that

$$\sum_z p(z) \inf_{\zeta^z \in \text{LHS}} D(\omega_{X_f \hat{A}_f B'_f Y}^z \|\zeta_{X_f \hat{A}_f B'_f Y}^z) \leq \inf_{\{\hat{\sigma}_B^{a, x}\}_{a, x} \in \text{LHS}} D(\omega_{X_f X \hat{A}_{B'_f} Y Z} \|\tau_{X_f X \hat{A}_{B'_f} Y Z}). \quad (3.33)$$

The above inequality holds for all 1W-LOCC strategies $\{p_{X_f|YZ=z}, \{L_y^{(z)}\}_y\}_z$, so we can now take a supremum over all such strategies $\{p_{X_f|YZ=z}, \{L_y^{(z)}\}_y\}_z$ to find that

$$\begin{aligned} & \sup_{\{p_{X_f|YZ=z}, \{L_y^{(z)}\}_y\}_z} \sum_z p(z) \inf_{\zeta^z \in \text{LHS}} D(\omega_{X_f \hat{A}_f B'_f Y}^z \|\zeta_{X_f \hat{A}_f B'_f Y}^z) \\ & \leq \sup_{\{p_{X_f|YZ=z}, \{L_y^{(z)}\}_y\}_z} \inf_{\{\hat{\sigma}_B^{a, x}\}_{a, x} \in \text{LHS}} D(\omega_{X_f X \hat{A}_{B'_f} Y Z} \|\tau_{X_f X \hat{A}_{B'_f} Y Z}) \end{aligned} \quad (3.34)$$

$$\leq R_S(\hat{A}; B)_{\hat{\rho}}. \quad (3.35)$$

The last inequality follows because $\{p_{X_f|YZ=z}, \{L_y^{(z)}\}_y\}_z$ is a particular 1W-LOCC strategy, while $R_S(\hat{A}; B)_{\hat{\rho}}$ involves an optimization over all 1W-LOCC strategies. The quantity on the first line above can be rewritten as

$$\begin{aligned} & \sup_{\{p_{X_f|YZ=z}, \{L_y^{(z)}\}_y\}_z} \sum_z p(z) \inf_{\zeta^z \in \text{LHS}} D(\omega_{X_f \hat{A}_f B'_f Y}^z \|\zeta_{X_f \hat{A}_f B'_f Y}^z) \\ & = \sum_z p(z) \sup_{\{p_{X_f|YZ=z}, \{L_y^{(z)}\}_y\}} \inf_{\zeta^z \in \text{LHS}} D(\omega_{X_f \hat{A}_f B'_f Y}^z \|\zeta_{X_f \hat{A}_f B'_f Y}^z) \end{aligned} \quad (3.36)$$

$$= \sum_z p(z) R_S(\hat{A}_f; B_f)_{\hat{\rho}_z}. \quad (3.37)$$

This concludes the proof. \blacksquare

3.3 Convexity

Here we prove that the relative entropy of steering is convex with respect to the assemblages on which it is evaluated.

Proposition 5 (Convexity) *Let $\lambda \in [0, 1]$. Let $\{\hat{\rho}_B^{a, x}\}_{a, x}$ and $\{\hat{\theta}_B^{a, x}\}_{a, x}$ be two assemblages, and consider an assemblage $\{\hat{\tau}_B^{a, x} := \lambda \hat{\rho}_B^{a, x} + (1 - \lambda) \hat{\theta}_B^{a, x}\}_{a, x}$. The restricted relative entropy of steering*

is convex in the following sense:

$$R_S(\hat{A}; B)_{\hat{\tau}} \leq \lambda R_S(\hat{A}; B)_{\hat{\rho}} + (1 - \lambda) R_S(\hat{A}; B)_{\hat{\theta}}. \quad (3.38)$$

Proof. Let $\{p_{X|Y}, \{K_y\}_y\}$ denote an arbitrary 1W-LOCC operation, let $\{\hat{\sigma}_B^{a,x}\}_{a,x}$ and $\{\hat{\omega}_B^{a,x}\}_{a,x}$ be arbitrary LHS assemblages. Consider the following states:

$$\rho_{X\hat{A}B'Y} := \sum_{x,a,y} p_{X|Y}(x|y)[x] \otimes [a] \otimes K_y \hat{\rho}_B^{a,x} K_y^\dagger \otimes [y], \quad (3.39)$$

$$\theta_{X\hat{A}B'Y} := \sum_{x,a,y} p_{X|Y}(x|y)[x] \otimes [a] \otimes K_y \hat{\theta}_B^{a,x} K_y^\dagger \otimes [y], \quad (3.40)$$

$$\sigma_{X\hat{A}B'Y} := \sum_{x,a,y} p_{X|Y}(x|y)[x] \otimes [a] \otimes K_y \hat{\sigma}_B^{a,x} K_y^\dagger \otimes [y], \quad (3.41)$$

$$\omega_{X\hat{A}B'Y} := \sum_{x,a,y} p_{X|Y}(x|y)[x] \otimes [a] \otimes K_y \hat{\omega}_B^{a,x} K_y^\dagger \otimes [y]. \quad (3.42)$$

Let us define the following states:

$$\zeta_{QX\hat{A}B'Y} := \lambda |0\rangle\langle 0|_Q \otimes \rho_{X\hat{A}B'Y} + (1 - \lambda) |1\rangle\langle 1|_Q \otimes \theta_{X\hat{A}B'Y}, \quad (3.43)$$

$$\kappa_{QX\hat{A}B'Y} := \lambda |0\rangle\langle 0|_Q \otimes \sigma_{X\hat{A}B'Y} + (1 - \lambda) |1\rangle\langle 1|_Q \otimes \omega_{X\hat{A}B'Y}. \quad (3.44)$$

Consider that

$$\zeta_{X\hat{A}B'Y} = \text{Tr}_Q(\zeta_{QX\hat{A}B'Y}) = \sum_{x,a,y} p_{X|Y}(x|y)[x] \otimes [a] \otimes K_y \hat{\tau}_B^{a,x} K_y^\dagger \otimes [y]. \quad (3.45)$$

Then we have the following chain of inequalities:

$$\begin{aligned} & \lambda D(\rho_{X\hat{A}B'Y} \| \sigma_{X\hat{A}B'Y}) + (1 - \lambda) D(\theta_{X\hat{A}B'Y} \| \omega_{X\hat{A}B'Y}) \\ &= D(\zeta_{QX\hat{A}B'Y} \| \kappa_{QX\hat{A}B'Y}) \end{aligned} \quad (3.46)$$

$$\geq D(\zeta_{X\hat{A}B'Y} \| \kappa_{X\hat{A}B'Y}) \quad (3.47)$$

$$\geq \inf_{\hat{\varsigma} \in \text{LHS}} D(\zeta_{X\hat{A}B'Y} \| \varsigma_{X\hat{A}B'Y}). \quad (3.48)$$

In the first equality, we have exploited the property of quantum relative entropy in (2.7). The first inequality follows from the data processing inequality for quantum relative entropy, by tracing over system Q . The final inequality follows by defining the LHS assemblage $\hat{\varsigma} \equiv \{\hat{\varsigma}_B^{a,x}\}_{a,x}$, the corresponding state

$$\varsigma_{X\hat{A}B'Y} := \sum_{x,a,y} p_{X|Y}(x|y)[x] \otimes [a] \otimes K_y \hat{\varsigma}_B^{a,x} K_y^\dagger \otimes [y], \quad (3.49)$$

and taking an infimum with respect to all such LHS assemblages. Since we have shown that the above inequality holds for all LHS assemblages $\{\hat{\sigma}_B^{a,x}\}_{a,x}$ and $\{\hat{\omega}_B^{a,x}\}_{a,x}$, we can conclude that

$$\inf_{\hat{\varsigma} \in \text{LHS}} D(\zeta_{X\hat{A}B'Y} \| \varsigma_{X\hat{A}B'Y}) \leq \lambda \inf_{\hat{\sigma} \in \text{LHS}} D(\rho_{X\hat{A}B'Y} \| \sigma_{X\hat{A}B'Y}) + (1 - \lambda) \inf_{\hat{\omega} \in \text{LHS}} D(\theta_{X\hat{A}B'Y} \| \omega_{X\hat{A}B'Y}). \quad (3.50)$$

Finally, since we have shown the above inequality holds for an arbitrary 1W-LOCC operation $\{p_{X|Y}, \{K_y\}_y\}$, we can conclude that

$$\begin{aligned}
& \sup_{\{p_{X|Y}, \{K_y\}_y\}} \inf_{\zeta \in \text{LHS}} D(\zeta_{X\hat{A}B'Y} \| \varsigma_{X\hat{A}B'Y}) \\
& \leq \sup_{\{p_{X|Y}, \{K_y\}_y\}} \left[\lambda \inf_{\hat{\sigma} \in \text{LHS}} D(\rho_{X\hat{A}B'Y} \| \sigma_{X\hat{A}B'Y}) + (1 - \lambda) \inf_{\hat{\omega} \in \text{LHS}} D(\theta_{X\hat{A}B'Y} \| \omega_{X\hat{A}B'Y}) \right] \quad (3.51) \\
& \leq \lambda \sup_{\{p_{X|Y}, \{K_y\}_y\}} \inf_{\hat{\sigma} \in \text{LHS}} D(\rho_{X\hat{A}B'Y} \| \sigma_{X\hat{A}B'Y}) + (1 - \lambda) \sup_{\{p_{X|Y}, \{K_y\}_y\}} \inf_{\hat{\omega} \in \text{LHS}} D(\theta_{X\hat{A}B'Y} \| \omega_{X\hat{A}B'Y}). \quad (3.52)
\end{aligned}$$

This final inequality is equivalent to the one in the statement of the proposition. ■

3.4 Upper bounds on relative entropy of steering

Proposition 6 (Upper bounds) *Let $\{\hat{\rho}_B^{a,x}\}_{a,x}$ be an assemblage. Then*

$$R_S(\hat{A}; B)_{\hat{\rho}} \leq \sup_{\{p_{X|Y}, \{K_y\}_y\}} I(XB'Y; \hat{A})_{\rho} \leq \sup_{p_X} H(\hat{A}) \leq \log_2 |\hat{A}|, \quad (3.53)$$

where the mutual information is with respect to the following state:

$$\rho_{X\hat{A}B'Y} := \sum_{x,a,y} p_{X|Y}(x|y)[x] \otimes [a] \otimes K_y \hat{\rho}_B^{a,x} K_y^\dagger \otimes [y], \quad (3.54)$$

and the entropy $H(\hat{A})$ is with respect to the probability distribution $p_{\hat{A}}(a) := \sum_x p_X(x) \text{Tr}(\hat{\rho}_B^{a,x})$.

Proof. From the definition of relative entropy of steering, we have that

$$R_S(\hat{A}; B)_{\hat{\rho}} = \sup_{\{p_{X|Y}, \{K_y\}_y\}} \inf_{\{\hat{\sigma}_B^{a,x}\}_{a,x} \in \text{LHS}} D(\rho_{X\hat{A}B'Y} \| \sigma_{X\hat{A}B'Y}), \quad (3.55)$$

where

$$\sigma_{X\hat{A}B'Y} := \sum_{x,a,y} p_{X|Y}(x|y)[x] \otimes [a] \otimes K_y \hat{\sigma}_B^{a,x} K_y^\dagger \otimes [y]. \quad (3.56)$$

Consider the probability distribution on system \hat{A} that results from partial trace with respect to the state $\rho_{X\hat{A}B'Y}$:

$$\rho_{\hat{A}} = \text{Tr}_{XB'Y}(\rho_{X\hat{A}B'Y}) \quad (3.57)$$

$$= \sum_a \left[\sum_{x,y} p_{X|Y}(x|y) \text{Tr}(K_y \hat{\rho}_B^{a,x} K_y^\dagger) \right] [a]. \quad (3.58)$$

Then define $p_A(a) := \sum_{x,y} p_{X|Y}(x|y) \text{Tr}(K_y \hat{\rho}_B^{a,x} K_y^\dagger)$. Also, take $\rho_B := \sum_a \hat{\rho}_B^{a,x}$. A particular assemblage with a local-hidden-state model is the following one:

$$\{\xi_B^{a,x} := p_A(a) \rho_B\}_{a,x}. \quad (3.59)$$

This particular LHS assemblage leads to the following state on systems $X\hat{A}B'Y$:

$$\sum_{x,a,y} p_{X|Y}(x|y)[x] \otimes p_A(a)[a] \otimes K_y \rho_B K_y^\dagger \otimes [y] = \rho_{\hat{A}} \otimes \rho_{XB'Y}, \quad (3.60)$$

where the states on the right are the marginals of $\rho_{X\hat{A}B'Y}$. Then

$$\begin{aligned} & \sup_{\{p_{X|Y}, \{K_y\}_y\}} \inf_{\{\hat{\rho}_B^{a,x}\}_{a,x} \in \text{LHS}} D(\rho_{X\hat{A}B'Y} \| \sigma_{X\hat{A}B'Y}) \\ & \leq \sup_{\{p_{X|Y}, \{K_y\}_y\}} D(\rho_{X\hat{A}B'Y} \| \rho_{\hat{A}} \otimes \rho_{XB'Y}) \end{aligned} \quad (3.61)$$

$$= \sup_{\{p_{X|Y}, \{K_y\}_y\}} I(XB'Y; \hat{A})_\rho \quad (3.62)$$

$$\leq \sup_{p_X} H(\hat{A})_\rho \quad (3.63)$$

$$\leq \log_2 |\hat{A}|. \quad (3.64)$$

The first inequality follows because we can choose a particular LHS assemblage and get an upper bound on $R_S(\hat{A}; B)_{\hat{\rho}}$. The first equality follows from the well known characterization of quantum mutual information as the quantum relative entropy between the joint state and the product of the marginals. The second inequality follows because $I(XB'Y; \hat{A})_\rho \leq H(\hat{A})_\rho$, given that system \hat{A} is classical, and then we can optimize this quantity with respect to all possible input distributions p_X . The final inequality is a well known dimension bound for entropy. ■

3.5 Continuity

Before giving the statement of continuity, let us first define the (normalized) trace distance of assemblages as follows:

Definition 7 (Trace distance of assemblages) Let $\{\hat{\rho}_B^{a,x}\}_{a,x}$ and $\{\hat{\theta}_B^{a,x}\}_{a,x}$ be two assemblages. We define the normalized trace distance of assemblages as

$$\Delta(\hat{\rho}, \hat{\theta}) := \frac{1}{2} \sup_{\{p_{X|Y}, \{K_y\}_y\}} \|\rho_{X\hat{A}B'Y} - \theta_{X\hat{A}B'Y}\|_1, \quad (3.65)$$

where $\|C\|_1 := \text{Tr}(\sqrt{C^\dagger C})$ and

$$\rho_{X\hat{A}B'Y} := \sum_{x,a,y} p_{X|Y}(x|y)[x] \otimes [a] \otimes K_y \hat{\rho}_B^{a,x} K_y^\dagger \otimes [y], \quad (3.66)$$

$$\theta_{X\hat{A}B'Y} := \sum_{x,a,y} p_{X|Y}(x|y)[x] \otimes [a] \otimes K_y \hat{\theta}_B^{a,x} K_y^\dagger \otimes [y]. \quad (3.67)$$

By properties of trace distance, it follows that $\Delta(\hat{\rho}, \hat{\theta}) \in [0, 1]$.

Proposition 8 (Metric) The trace distance of assemblages is a metric, in the sense that for any

three assemblages $\{\hat{\rho}_B^{a,x}\}_{a,x}$, $\{\hat{\theta}_B^{a,x}\}_{a,x}$, and $\{\hat{\omega}_B^{a,x}\}_{a,x}$:

$$\Delta(\hat{\rho}, \hat{\theta}) \geq 0, \quad (3.68)$$

$$\Delta(\hat{\rho}, \hat{\theta}) = 0 \text{ if and only if } \hat{\rho}_B^{a,x} = \hat{\theta}_B^{a,x} \text{ for all } a, x, \quad (3.69)$$

$$\Delta(\hat{\rho}, \hat{\theta}) = \Delta(\hat{\theta}, \hat{\rho}), \quad (3.70)$$

$$\Delta(\hat{\rho}, \hat{\theta}) \leq \Delta(\hat{\rho}, \hat{\omega}) + \Delta(\hat{\omega}, \hat{\theta}). \quad (3.71)$$

Proof. These properties follow directly from the fact that normalized trace distance is a metric for quantum states. We give brief proofs for completeness. The inequality in (3.68) follows because the normalized trace distance is non-negative. Regarding (3.69), the implication $\hat{\rho}_B^{a,x} = \hat{\theta}_B^{a,x}$ for all $a, x \implies \Delta(\hat{\rho}, \hat{\theta}) = 0$ follows because the states resulting from an arbitrary 1W-LOCC operation are the same if the assemblages are the same. To see the other implication, consider that $\Delta(\hat{\rho}, \hat{\theta}) = 0$ means that the normalized trace distance between $\rho_{X\hat{A}B'Y}$ and $\theta_{X\hat{A}B'Y}$ is equal to zero for all possible 1W-LOCC operations. So we can pick the 1W-LOCC operation to be a uniform distribution over the input x and the identity channel on system B and find that

$$0 = \left\| \sum_{x,a} \frac{1}{|\mathcal{X}|} [x] \otimes [a] \otimes \hat{\rho}_B^{a,x} - \sum_{x,a} \frac{1}{|\mathcal{X}|} [x] \otimes [a] \otimes \hat{\theta}_B^{a,x} \right\|_1 \quad (3.72)$$

$$= \sum_{x,a} \frac{1}{|\mathcal{X}|} \left\| \hat{\rho}_B^{a,x} - \hat{\theta}_B^{a,x} \right\|_1. \quad (3.73)$$

By the fact that the normalized trace distance is a metric, we can then conclude that $\hat{\rho}_B^{a,x} = \hat{\theta}_B^{a,x}$ for all a, x . The equality in (3.70) clearly holds. The triangle inequality in (3.71) follows because normalized trace distance obeys the triangle inequality:

$$\Delta(\hat{\rho}, \hat{\theta}) = \frac{1}{2} \sup_{\{p_{X|Y}, \{K_y\}_y\}} \left\| \rho_{X\hat{A}B'Y} - \theta_{X\hat{A}B'Y} \right\|_1 \quad (3.74)$$

$$\leq \frac{1}{2} \sup_{\{p_{X|Y}, \{K_y\}_y\}} \left[\left\| \rho_{X\hat{A}B'Y} - \omega_{X\hat{A}B'Y} \right\|_1 + \left\| \omega_{X\hat{A}B'Y} - \theta_{X\hat{A}B'Y} \right\|_1 \right] \quad (3.75)$$

$$\leq \frac{1}{2} \sup_{\{p_{X|Y}, \{K_y\}_y\}} \left\| \rho_{X\hat{A}B'Y} - \omega_{X\hat{A}B'Y} \right\|_1 + \frac{1}{2} \sup_{\{p_{X|Y}, \{K_y\}_y\}} \left\| \omega_{X\hat{A}B'Y} - \theta_{X\hat{A}B'Y} \right\|_1 \quad (3.76)$$

$$= \Delta(\hat{\rho}, \hat{\omega}) + \Delta(\hat{\omega}, \hat{\theta}). \quad (3.77)$$

This concludes the proof. ■

Theorem 9 (Uniform continuity) Let $\{\hat{\rho}_B^{a,x}\}_{a,x}$ and $\{\hat{\theta}_B^{a,x}\}_{a,x}$ be two assemblages such that $\Delta(\hat{\rho}, \hat{\theta}) \leq \varepsilon \in [0, 1]$. Then

$$\left| R_S(\hat{A}; B)_{\hat{\rho}} - R_S(\hat{A}; B)_{\hat{\theta}} \right| \leq \varepsilon \log_2 |\hat{A}| + g(\varepsilon), \quad (3.78)$$

where $g(\varepsilon) := (\varepsilon + 1) \log_2(\varepsilon + 1) - \varepsilon \log_2 \varepsilon$.

Proof. We note that the following proof is very similar to those of [Win16, Lemmas 2 and 7], but we give a detailed proof for completeness. Let $\{p_{X|Y}, \{K_y\}_y\}$ be an arbitrary 1W-LOCC operation,

and let $\{\hat{\sigma}_B^{a,x}\}_{a,x}$ and $\{\hat{\omega}_B^{a,x}\}_{a,x}$ be arbitrary LHS assemblages. Consider the following states:

$$\rho_{X\hat{A}B'Y} := \sum_{x,a,y} p_{X|Y}(x|y)[x] \otimes [a] \otimes K_y \hat{\rho}_B^{a,x} K_y^\dagger \otimes [y], \quad (3.79)$$

$$\theta_{X\hat{A}B'Y} := \sum_{x,a,y} p_{X|Y}(x|y)[x] \otimes [a] \otimes K_y \hat{\theta}_B^{a,x} K_y^\dagger \otimes [y], \quad (3.80)$$

$$\sigma_{X\hat{A}B'Y} := \sum_{x,a,y} p_{X|Y}(x|y)[x] \otimes [a] \otimes K_y \hat{\sigma}_B^{a,x} K_y^\dagger \otimes [y], \quad (3.81)$$

$$\omega_{X\hat{A}B'Y} := \sum_{x,a,y} p_{X|Y}(x|y)[x] \otimes [a] \otimes K_y \hat{\omega}_B^{a,x} K_y^\dagger \otimes [y]. \quad (3.82)$$

Consider that $\frac{1}{2} \|\rho_{X\hat{A}B'Y} - \theta_{X\hat{A}B'Y}\|_1 \leq \varepsilon$ by assumption. Let us set $\varepsilon_0 := \frac{1}{2} \|\rho_{X\hat{A}B'Y} - \theta_{X\hat{A}B'Y}\|_1$. If $\varepsilon_0 = 0$, then the particular 1W-LOCC operation cannot distinguish the states from each other, so that $\rho_{X\hat{A}B'Y} = \theta_{X\hat{A}B'Y}$, and we find that

$$\inf_{\hat{\sigma} \in \text{LHS}} D(\rho_{X\hat{A}B'Y} \| \sigma_{X\hat{A}B'Y}) = \inf_{\hat{\theta} \in \text{LHS}} D(\theta_{X\hat{A}B'Y} \| \sigma_{X\hat{A}B'Y}) \quad (3.83)$$

in this case, so that there is nothing to prove. So let us instead suppose that $\varepsilon_0 \neq 0$ and define

$$\Delta_{X\hat{A}B'Y} := \frac{1}{\varepsilon_0} (\rho_{X\hat{A}B'Y} - \theta_{X\hat{A}B'Y})_+, \quad (3.84)$$

where $(\cdot)_+$ indicates the positive part of $\rho_{X\hat{A}B'Y} - \theta_{X\hat{A}B'Y}$. Since $\rho_{X\hat{A}B'Y} - \theta_{X\hat{A}B'Y}$ is traceless and its trace norm is equal to $2\varepsilon_0$, it follows that $\Delta_{X\hat{A}B'Y}$ is a density operator. Consider that

$$\rho_{X\hat{A}B'Y} = \theta_{X\hat{A}B'Y} + (\rho_{X\hat{A}B'Y} - \theta_{X\hat{A}B'Y}) \quad (3.85)$$

$$\leq \theta_{X\hat{A}B'Y} + \varepsilon_0 \Delta_{X\hat{A}B'Y} \quad (3.86)$$

$$= (1 + \varepsilon_0) \left(\frac{1}{1 + \varepsilon_0} \theta_{X\hat{A}B'Y} + \frac{\varepsilon_0}{1 + \varepsilon_0} \Delta_{X\hat{A}B'Y} \right) \quad (3.87)$$

$$=: (1 + \varepsilon_0) \zeta_{X\hat{A}B'Y}. \quad (3.88)$$

Setting

$$\Delta'_{X\hat{A}B'Y} := \frac{1}{\varepsilon_0} [(1 + \varepsilon_0) \zeta_{X\hat{A}B'Y} - \rho_{X\hat{A}B'Y}], \quad (3.89)$$

we see that $\Delta'_{X\hat{A}B'Y}$ is a density operator also, satisfying

$$\zeta_{X\hat{A}B'Y} = \frac{1}{1 + \varepsilon_0} \theta_{X\hat{A}B'Y} + \frac{\varepsilon_0}{1 + \varepsilon_0} \Delta_{X\hat{A}B'Y} \quad (3.90)$$

$$= \frac{1}{1 + \varepsilon_0} \rho_{X\hat{A}B'Y} + \frac{\varepsilon_0}{1 + \varepsilon_0} \Delta'_{X\hat{A}B'Y}. \quad (3.91)$$

From the joint convexity of relative entropy, we find that

$$\begin{aligned} & \inf_{\hat{\kappa} \in \text{LHS}} D(\zeta_{X\hat{A}B'Y} \| \kappa_{X\hat{A}B'Y}) \\ & \leq D(\zeta_{X\hat{A}B'Y} \| [1 + \varepsilon_0]^{-1} \sigma_{X\hat{A}B'Y} + \varepsilon_0 [1 + \varepsilon_0]^{-1} \omega_{X\hat{A}B'Y}) \end{aligned} \quad (3.92)$$

$$\leq \frac{1}{1 + \varepsilon_0} D(\theta_{X\hat{A}B'Y} \| \sigma_{X\hat{A}B'Y}) + \frac{\varepsilon_0}{1 + \varepsilon_0} D(\Delta_{X\hat{A}B'Y} \| \omega_{X\hat{A}B'Y}). \quad (3.93)$$

Since $\sigma_{X\hat{A}B'Y}$ and $\omega_{X\hat{A}B'Y}$ are states arising from arbitrary LHS assemblages, we can conclude that

$$\inf_{\hat{\kappa} \in \text{LHS}} D(\zeta_{X\hat{A}B'Y} \| \kappa_{X\hat{A}B'Y}) \leq \frac{1}{1 + \varepsilon_0} \inf_{\hat{\sigma} \in \text{LHS}} D(\theta_{X\hat{A}B'Y} \| \sigma_{X\hat{A}B'Y}) + \frac{\varepsilon_0}{1 + \varepsilon_0} \inf_{\hat{\omega} \in \text{LHS}} D(\Delta_{X\hat{A}B'Y} \| \omega_{X\hat{A}B'Y}). \quad (3.94)$$

Now consider that for a state $\kappa_{X\hat{A}B'Y}$ arising from an arbitrary LHS assemblage $\hat{\kappa}$, we have that

$$\begin{aligned} D(\zeta_{X\hat{A}B'Y} \| \kappa_{X\hat{A}B'Y}) &= -H(\zeta_{X\hat{A}B'Y}) - \text{Tr}(\zeta_{X\hat{A}B'Y} \log_2 \kappa_{X\hat{A}B'Y}) \end{aligned} \quad (3.95)$$

$$\begin{aligned} &\geq -h_2(\varepsilon_0/[1 + \varepsilon_0]^{-1}) - \frac{1}{1 + \varepsilon_0} H(\rho_{X\hat{A}B'Y}) - \frac{\varepsilon_0}{1 + \varepsilon_0} H(\Delta'_{X\hat{A}B'Y}) \\ &\quad - \frac{1}{1 + \varepsilon_0} \text{Tr}(\rho_{X\hat{A}B'Y} \log \kappa_{X\hat{A}B'Y}) - \frac{\varepsilon_0}{1 + \varepsilon_0} \text{Tr}(\Delta'_{X\hat{A}B'Y} \log \kappa_{X\hat{A}B'Y}) \end{aligned} \quad (3.96)$$

$$\begin{aligned} &= -h_2(\varepsilon_0/[1 + \varepsilon_0]^{-1}) + \frac{1}{1 + \varepsilon_0} D(\rho_{X\hat{A}B'Y} \| \kappa_{X\hat{A}B'Y}) \\ &\quad + \frac{\varepsilon_0}{1 + \varepsilon_0} D(\Delta'_{X\hat{A}B'Y} \| \kappa_{X\hat{A}B'Y}) \end{aligned} \quad (3.97)$$

$$\begin{aligned} &\geq -h_2(\varepsilon_0/[1 + \varepsilon_0]^{-1}) + \frac{1}{1 + \varepsilon_0} \inf_{\hat{\sigma} \in \text{LHS}} D(\rho_{X\hat{A}B'Y} \| \sigma_{X\hat{A}B'Y}) \\ &\quad + \frac{\varepsilon_0}{1 + \varepsilon_0} \inf_{\hat{\omega} \in \text{LHS}} D(\Delta'_{X\hat{A}B'Y} \| \omega_{X\hat{A}B'Y}). \end{aligned} \quad (3.98)$$

The first inequality follows because

$$H(\lambda \xi_0 + (1 - \lambda) \xi_1) \leq H(\{\lambda, 1 - \lambda\}) + \lambda H(\xi_0) + (1 - \lambda) H(\xi_1) \quad (3.99)$$

for $\lambda \in [0, 1]$ and density operators ξ_0 and ξ_1 and where we define $h_2(\lambda) := H(\{\lambda, 1 - \lambda\})$. Since we have shown that the above inequality holds for an arbitrary state $\kappa_{X\hat{A}B'Y}$ arising from an LHS assemblage $\hat{\kappa}$, we can conclude that

$$\begin{aligned} \inf_{\hat{\kappa} \in \text{LHS}} D(\zeta_{X\hat{A}B'Y} \| \kappa_{X\hat{A}B'Y}) &\geq -h_2(\varepsilon_0/[1 + \varepsilon_0]^{-1}) \\ &\quad + \frac{1}{1 + \varepsilon_0} \inf_{\hat{\sigma} \in \text{LHS}} D(\rho_{X\hat{A}B'Y} \| \sigma_{X\hat{A}B'Y}) + \frac{\varepsilon_0}{1 + \varepsilon_0} \inf_{\hat{\omega} \in \text{LHS}} D(\Delta'_{X\hat{A}B'Y} \| \omega_{X\hat{A}B'Y}). \end{aligned} \quad (3.100)$$

Putting the bounds in (3.94) and (3.100) together and multiplying by $1 + \varepsilon_0$, we conclude that

$$\begin{aligned} &\inf_{\hat{\sigma} \in \text{LHS}} D(\rho_{X\hat{A}B'Y} \| \sigma_{X\hat{A}B'Y}) + \varepsilon_0 \inf_{\hat{\omega} \in \text{LHS}} D(\Delta'_{X\hat{A}B'Y} \| \omega_{X\hat{A}B'Y}) - g(\varepsilon_0) \\ &\leq \inf_{\hat{\sigma} \in \text{LHS}} D(\theta_{X\hat{A}B'Y} \| \sigma_{X\hat{A}B'Y}) + \varepsilon_0 \inf_{\hat{\omega} \in \text{LHS}} D(\Delta_{X\hat{A}B'Y} \| \omega_{X\hat{A}B'Y}) \end{aligned} \quad (3.101)$$

$$\leq \inf_{\hat{\sigma} \in \text{LHS}} D(\theta_{X\hat{A}B'Y} \| \sigma_{X\hat{A}B'Y}) + \varepsilon_0 \log_2 |\hat{A}|, \quad (3.102)$$

where we have used that $g(\varepsilon_0) = (1 + \varepsilon_0) h_2(\varepsilon_0/[1 + \varepsilon_0]^{-1})$ [Shi15] and Proposition 6. By dropping the term $\varepsilon_0 \inf_{\hat{\omega} \in \text{LHS}} D(\Delta'_{X\hat{A}B'Y} \| \omega_{X\hat{A}B'Y})$ (it is non-negative), we can rewrite the above bound as

$$\inf_{\hat{\sigma} \in \text{LHS}} D(\rho_{X\hat{A}B'Y} \| \sigma_{X\hat{A}B'Y}) \leq \inf_{\hat{\sigma} \in \text{LHS}} D(\theta_{X\hat{A}B'Y} \| \sigma_{X\hat{A}B'Y}) + \varepsilon_0 \log_2 |\hat{A}| + g(\varepsilon_0) \quad (3.103)$$

$$\leq \inf_{\hat{\sigma} \in \text{LHS}} D(\theta_{X\hat{A}B'Y} \| \sigma_{X\hat{A}B'Y}) + \varepsilon \log_2 |\hat{A}| + g(\varepsilon), \quad (3.104)$$

where in the last line we have used the facts that $\varepsilon_0 \leq \varepsilon$ and the function $\varepsilon \log_2 |\hat{A}| + g(\varepsilon)$ is monotone non-decreasing with respect to ε . Since the above inequality holds for an arbitrary 1W-LOCC operation $\{p_{X|Y}, \{K_y\}_y\}$, we can conclude that

$$\begin{aligned} \sup_{\{p_{X|Y}, \{K_y\}_y\}} \inf_{\hat{\sigma} \in \text{LHS}} D(\rho_{X\hat{A}B'Y} \| \sigma_{X\hat{A}B'Y}) \\ \leq \sup_{\{p_{X|Y}, \{K_y\}_y\}} \inf_{\hat{\sigma} \in \text{LHS}} D(\theta_{X\hat{A}B'Y} \| \sigma_{X\hat{A}B'Y}) + \varepsilon \log_2 |\hat{A}| + g(\varepsilon), \end{aligned} \quad (3.105)$$

which is the same as

$$R_S(\hat{A}; B)_{\hat{\rho}} \leq R_S(\hat{A}; B)_{\hat{\theta}} + \varepsilon \log_2 |\hat{A}| + g(\varepsilon). \quad (3.106)$$

To get the other inequality $R_S(\hat{A}; B)_{\hat{\theta}} \leq R_S(\hat{A}; B)_{\hat{\rho}} + \varepsilon \log_2 |\hat{A}| + g(\varepsilon)$, we simply repeat all of the above steps with $\hat{\rho}$ and $\hat{\theta}$ swapped. ■

3.6 Faithfulness

A steering quantifier is *faithful* if it is equal to zero if and only if the assemblage has a local-hidden-state model. In this section, we prove quantitative statements regarding the faithfulness of relative entropy of steering. We begin with the implication $\hat{\rho} \in \text{LHS} \implies R_S(\hat{A}; B)_{\hat{\rho}} = 0$.

Proposition 10 *Let $\varepsilon \in [0, 1]$, and let $\{\hat{\rho}_B^{a,x}\}_{a,x}$ and $\{\hat{\sigma}_B^{a,x}\}_{a,x}$ be assemblages such that $\hat{\sigma} \in \text{LHS}$ and $\Delta(\hat{\rho}, \hat{\sigma}) \leq \varepsilon$. Then*

$$R_S(\hat{A}; B)_{\hat{\rho}} \leq \varepsilon \log_2 |\hat{A}| + g(\varepsilon). \quad (3.107)$$

Proof. This is a direct consequence of Proposition 9 and the fact that $\hat{\sigma} \in \text{LHS}$ so that by applying Definition 1, we see that $R_S(\hat{A}; B)_{\hat{\sigma}} = 0$. ■

We now establish the implication $R_S(\hat{A}; B)_{\hat{\rho}} = 0 \implies \hat{\rho} \in \text{LHS}$:

Proposition 11 *Let $\{\hat{\rho}_B^{a,x}\}_{a,x}$ be an assemblage. Then*

$$\sqrt{2 \ln 2} R_S(\hat{A}; B)_{\hat{\rho}} \geq \inf_{\{\hat{\sigma}_B^{a,x}\}_{a,x} \in \text{LHS}} \frac{1}{|\mathcal{X}|} \sum_{x,a} \|\hat{\rho}_B^{a,x} - \hat{\sigma}_B^{a,x}\|_1. \quad (3.108)$$

In particular, if $R_S(\hat{A}; B)_{\hat{\rho}} = 0$, then $\{\hat{\rho}_B^{a,x}\}_{a,x} \in \text{LHS}$.

Proof. The inequality in (3.108) is a direct consequence of the quantum Pinsker inequality [OP93, Theorem 1.15], which is the statement that

$$D(\omega \| \tau) \geq \frac{1}{2 \ln 2} \|\omega - \tau\|_1^2, \quad (3.109)$$

for quantum states ω and τ . Applying it and definitions, we find that

$$\begin{aligned} & \sqrt{2 \ln 2} R_S(\hat{A}; B)_{\hat{\rho}} \\ & \geq \sup_{\{p_{X|Y}, \{K_y\}_y\}} \inf_{\{\hat{\sigma}_B^{a,x}\}_{a,x} \in \text{LHS}} \|\rho_{XAB'Y} - \sigma_{XAB'Y}\|_1 \end{aligned} \quad (3.110)$$

$$\geq \inf_{\{\hat{\sigma}_B^{a,x}\}_{a,x} \in \text{LHS}} \left\| \frac{1}{|\mathcal{X}|} \sum_{x,a} [x] \otimes [a] \otimes \hat{\rho}_B^{a,x} - \frac{1}{|\mathcal{X}|} \sum_{x,a} [x] \otimes [a] \otimes \hat{\sigma}_B^{a,x} \right\|_1 \quad (3.111)$$

$$= \inf_{\{\hat{\sigma}_B^{a,x}\}_{a,x} \in \text{LHS}} \frac{1}{|\mathcal{X}|} \sum_{x,a} \|\hat{\rho}_B^{a,x} - \hat{\sigma}_B^{a,x}\|_1, \quad (3.112)$$

where the second inequality follows by picking a 1W-LOCC operation to be trivial, consisting of choosing the input x uniformly at random and applying the identity channel to system B .

To get the implication $R_S(\hat{A}; B)_{\hat{\rho}} = 0 \implies \{\hat{\rho}_B^{a,x}\}_{a,x} \in \text{LHS}$, consider that the trace distance is continuous and the set LHS is compact, so that the infimum can be replaced with a minimum and thus in the case that $R_S(\hat{A}; B)_{\hat{\rho}} = 0$, we can conclude that there exists $\{\hat{\sigma}_B^{a,x}\}_{a,x} \in \text{LHS}$ such that $\hat{\rho}_B^{a,x} = \hat{\sigma}_B^{a,x}$ for all a and x . ■

4 Restricted relative entropy of steering

We also have a restricted relative entropy of steering, defined as

Definition 12 (Restricted relative entropy of steering) *Let $\{\hat{\rho}_B^{a,x}\}_{a,x}$ be an assemblage. Then the restricted relative entropy of steering is given by*

$$R_S^R(\hat{A}; B)_{\hat{\rho}} := \sup_{p_X} \inf_{\{\hat{\sigma}_B^{a,x}\}_{a,x} \in \text{LHS}} D(\rho_{X\hat{A}B} \| \sigma_{X\hat{A}B}), \quad (4.1)$$

where

$$\rho_{X\hat{A}B} := \sum_{x,a} p_X(x) |x\rangle\langle x|_X \otimes |a\rangle\langle a|_{\hat{A}} \otimes \hat{\rho}_B^{a,x}, \quad (4.2)$$

$$\sigma_{X\hat{A}B} := \sum_{x,a} p_X(x) |x\rangle\langle x|_X \otimes |a\rangle\langle a|_{\hat{A}} \otimes \hat{\sigma}_B^{a,x}. \quad (4.3)$$

The restricted relative entropy of steering obeys many properties similar to those of the relative entropy of steering, and we mostly list them below without proof because their proofs follow quite similarly to what we have shown previously (i.e., in some cases, a proof seems necessary and so we give it, while in others, a proof is an immediate consequence of prior developments and so we do not give it).

The first is the following:

Theorem 13 (Restricted 1W-LOCC monotone) *Let $\{\hat{\rho}_B^{a,x}\}_{a,x}$ be an assemblage, and let*

$$\{p_{X|X_f}, p_{\hat{A}_f|\hat{A}X_fZ}, \{K_z\}_z\} \quad (4.4)$$

denote a restricted 1W-LOCC operation that results in an assemblage $\{\hat{\omega}_{B'}^{a_f, x_f}\}_{a_f, x_f}$, defined as

$$\hat{\omega}_{B'}^{a_f, x_f} := \sum_{a, x, z} p_{X|X_f}(x|x_f) p_{\hat{A}_f|\hat{A}X X_f Z}(a_f|a, x, x_f, z) K_z \hat{\rho}_B^{a, x} K_z^\dagger. \quad (4.5)$$

Then

$$R_S^R(\hat{A}; B)_{\hat{\rho}} \geq R_S^R(\hat{A}_f; B')_{\hat{\omega}}. \quad (4.6)$$

Proof. Taking a distribution p_{X_f} over the black-box inputs of the final assemblage, we can embed the state of the final assemblage into the following classical–quantum state:

$$\omega_{X_f \hat{A}_f B'} := \sum_{x_f, a_f} p_{X_f}(x_f) [x_f] \otimes [a_f] \otimes \hat{\omega}_{B'}^{a_f, x_f}, \quad (4.7)$$

which is a marginal of the following state:

$$\begin{aligned} \omega_{X_f X \hat{A}_f \hat{A} Z B'} &:= \sum_{x_f, a_f, a, x, z} p_{X_f}(x_f) [x_f] \otimes p_{X|X_f}(x|x_f) [x] \\ &\quad \otimes p_{\hat{A}_f|\hat{A}X X_f Z}(a_f|a, x, x_f, z) [a_f] \otimes [a] \otimes [z] \otimes K_z \hat{\rho}_B^{a, x} K_z^\dagger. \end{aligned} \quad (4.8)$$

Let $\hat{\sigma}_{B'}^{a_f, x_f}$ denote an arbitrary LHS assemblage, and let $\sigma_{X_f \hat{A}_f B'}$ denote its corresponding classical–quantum state:

$$\sigma_{X_f \hat{A}_f B'} := \sum_{x_f, a_f} p_{X_f}(x_f) [x_f] \otimes [a_f] \otimes \hat{\sigma}_{B'}^{a_f, x_f}. \quad (4.9)$$

Let $\hat{\tau}_B^{a, x}$ denote an arbitrary LHS assemblage, and let $\tau_{X_f X \hat{A}_f \hat{A} Z B'}$ denote the following state:

$$\begin{aligned} \tau_{X_f X \hat{A}_f \hat{A} Z B'} &:= \sum_{x_f, a_f, a, x, z} p_{X_f}(x_f) [x_f] \otimes p_{X|X_f}(x|x_f) [x] \\ &\quad \otimes p_{\hat{A}_f|\hat{A}X X_f Z}(a_f|a, x, x_f, z) [a_f] \otimes [a] \otimes [z] \otimes K_z \hat{\tau}_B^{a, x} K_z^\dagger. \end{aligned} \quad (4.10)$$

Define the following states:

$$\omega'_{X_f X \hat{A} B} := \sum_{x_f, a_f, a, x, z} p_{X_f}(x_f) [x_f] \otimes p_{X|X_f}(x|x_f) [x] \otimes [a] \otimes \hat{\rho}_B^{a, x}, \quad (4.11)$$

$$\tau'_{X_f X \hat{A} B} := \sum_{x_f, a_f, a, x, z} p_{X_f}(x_f) [x_f] \otimes p_{X|X_f}(x|x_f) [x] \otimes [a] \otimes \hat{\tau}_B^{a, x}. \quad (4.12)$$

Consider that $\tau_{X_f \hat{A}_f B'}$ corresponds to an LHS assemblage by [GA15, Theorem 1]. Then

$$\inf_{\{\hat{\sigma}_{B'}^{a_f, x_f}\}_{a_f, x_f} \in \text{LHS}} D(\omega_{X_f \hat{A}_f B'} \| \sigma_{X_f \hat{A}_f B'}) \leq D(\omega_{X_f \hat{A}_f B'} \| \tau_{X_f \hat{A}_f B'}) \quad (4.13)$$

$$\leq D(\omega_{X_f X \hat{A}_f \hat{A} Z B'} \| \tau_{X_f X \hat{A}_f \hat{A} Z B'}) \quad (4.14)$$

$$= D(\omega_{X_f X \hat{A} Z B'} \| \tau_{X_f X \hat{A} Z B'}) \quad (4.15)$$

$$\leq D(\omega'_{X_f X \hat{A} B} \| \tau'_{X_f X \hat{A} B}) \quad (4.16)$$

$$= D(\omega'_{X \hat{A} B} \| \tau'_{X \hat{A} B}). \quad (4.17)$$

The first inequality follows because $\tau_{X_f \hat{A}_f B'}$ corresponds to a particular LHS assemblage. The second inequality follows from the data-processing inequality. The equality follows due to

$$D(\omega_{X_f X \hat{A}_f \hat{A} Z B'} \| \tau_{X_f X \hat{A}_f \hat{A} Z B'}) \leq D(\omega_{X_f X \hat{A} Z B'} \| \tau_{X_f X \hat{A} Z B'}), \quad (4.18)$$

which is a consequence of the fact that register \hat{A}_f results from processing the values in $X_f X \hat{A} Z$ according to $p_{\hat{A}_f | \hat{A} X X_f Z}$, while the opposite inequality

$$D(\omega_{X_f X \hat{A}_f \hat{A} Z B'} \| \tau_{X_f X \hat{A}_f \hat{A} Z B'}) \geq D(\omega_{X_f X \hat{A} Z B'} \| \tau_{X_f X \hat{A} Z B'}) \quad (4.19)$$

follows because partial trace over \hat{A}_f is a channel. The final inequality again follows from data processing: We get $\omega_{X_f X \hat{A} Z B'}$ from $\omega'_{X_f X \hat{A} B}$ and $\tau_{X_f X \hat{A} Z B'}$ from $\tau'_{X_f X \hat{A} B}$ by performing the quantum channel $(\cdot) \rightarrow \sum_z [z] \otimes K_z(\cdot) K_z^\dagger$ on system B . The final equality follows again from data processing: $D(\omega'_{X_f X \hat{A} B} \| \tau'_{X_f X \hat{A} B}) \geq D(\omega'_{X \hat{A} B} \| \tau'_{X \hat{A} B})$ because partial trace is a channel and $D(\omega'_{X_f X \hat{A} B} \| \tau'_{X_f X \hat{A} B}) \leq D(\omega'_{X \hat{A} B} \| \tau'_{X \hat{A} B})$ because we can apply the Bayes theorem to see that $p_{X_f} p_{X|X_f} = p_{X_f|X} p_X$ and thus X_f can be seen to arise from processing of X . Since we have shown that the inequality holds for an arbitrary LHS assemblage, we can conclude that

$$\inf_{\{\hat{\sigma}_{B'}^{a_f, x_f}\}_{a_f, x_f \in \text{LHS}}} D(\omega_{X_f \hat{A}_f B'} \| \sigma_{X_f \hat{A}_f B'}) \leq \inf_{\{\hat{\tau}_B^{a, x}\}_{a, x \in \text{LHS}}} D(\omega'_{X \hat{A} B} \| \tau'_{X \hat{A} B}) \quad (4.20)$$

$$\leq \sup_{p_X} \inf_{\{\hat{\tau}_B^{a, x}\}_{a, x \in \text{LHS}}} D(\omega'_{X \hat{A} B} \| \tau'_{X \hat{A} B}) \quad (4.21)$$

$$= R_S^R(\hat{A}; B)_{\hat{\rho}}. \quad (4.22)$$

Since the above holds for an arbitrary distribution p_{X_f} , we can conclude that

$$\sup_{p_{X_f}} \inf_{\{\hat{\sigma}_{B'}^{a_f, x_f}\}_{a_f, x_f \in \text{LHS}}} D(\omega_{X_f \hat{A}_f B'} \| \sigma_{X_f \hat{A}_f B'}) \leq R_S^R(\hat{A}; B)_{\hat{\rho}}, \quad (4.23)$$

which is equivalent to the statement of the theorem. ■

Proposition 14 (Convexity) *Let $\lambda \in [0, 1]$. Let $\{\hat{\rho}_B^{a, x}\}_{a, x}$ and $\{\hat{\theta}_B^{a, x}\}_{a, x}$ be two assemblages, and consider an assemblage $\{\hat{\tau}_B^{a, x} := \lambda \hat{\rho}_B^{a, x} + (1 - \lambda) \hat{\theta}_B^{a, x}\}_{a, x}$. The restricted relative entropy of steering is convex in the following sense:*

$$R_S^R(\hat{A}; B)_{\hat{\tau}} \leq \lambda R_S^R(\hat{A}; B)_{\hat{\rho}} + (1 - \lambda) R_S^R(\hat{A}; B)_{\hat{\theta}}. \quad (4.24)$$

The next proposition finds several upper bounds on the restricted relative entropy of steering, one of which is in terms of the conditional mutual information, defined for a tripartite state ς_{KLM} as $I(K; L|M)_\varsigma := H(KM)_\varsigma + H(LM)_\varsigma - H(M)_\varsigma - H(KLM)_\varsigma$.

Proposition 15 (Upper bounds) *Let $\{\hat{\rho}_B^{a, x}\}_{a, x}$ be an assemblage. Then*

$$R_S^R(\hat{A}; B)_{\hat{\rho}} \leq \sup_{p_X} I(\hat{A}; B|X)_\rho \quad (4.25)$$

$$\leq \min \left\{ \sup_{p_X} H(\hat{A}), H(B)_\rho \right\} \quad (4.26)$$

$$\leq \min \left\{ \log_2 |\hat{A}|, \log_2 |B| \right\}, \quad (4.27)$$

where the conditional mutual information is with respect to the following state:

$$\rho_{X\hat{A}B'Y} := \sum_{x,a} p_X(x)[x] \otimes [a] \otimes \hat{\rho}_B^{a,x}, \quad (4.28)$$

and the entropy $H(\hat{A})$ is with respect to the probability distribution $p_{\hat{A}}(a) := \sum_x p_X(x) \text{Tr}(\hat{\rho}_B^{a,x})$.

Proof. We can choose an assemblage having a local-hidden state model to be as follows:

$$\{\hat{\xi}_B^{a,x} := p_{\hat{A}|X}(a|x)\rho_B\}_{a,x}, \quad (4.29)$$

where $p_{\hat{A}|X}(a|x) = \text{Tr}(\hat{\rho}_B^{a,x})$ and $\rho_B = \sum_a \hat{\rho}_B^{a,x}$ (recall the no-signaling condition in (2.1)). Then define the following state:

$$\xi_{X\hat{A}B} := \sum_{x,a} p_X(x)[x] \otimes [a] \otimes \hat{\xi}_B^{a,x} = \left[\sum_{x,a} p_X(x)p_{\hat{A}|X}(a|x)[x] \otimes [a] \right] \otimes \rho_B = \rho_{X\hat{A}} \otimes \rho_B. \quad (4.30)$$

Consider that

$$\inf_{\{\hat{\sigma}_B^{a,x}\}_{a,x} \in \text{LHS}} D(\rho_{X\hat{A}B} \| \sigma_{X\hat{A}B}) \leq D(\rho_{X\hat{A}B} \| \xi_{X\hat{A}B}) \quad (4.31)$$

$$= I(X\hat{A}; B)_\rho \quad (4.32)$$

$$= I(\hat{A}; B|X)_\rho + I(X; B)_\rho \quad (4.33)$$

$$= I(\hat{A}; B|X)_\rho. \quad (4.34)$$

The inequality follows because the state $\xi_{X\hat{A}B}$ arises from a particular LHS assemblage. The first equality follows from the well known characterization of mutual information as the relative entropy of the joint state to the product of the marginals. The second equality follows from the chain rule for conditional mutual information (see, e.g., [Wil16]), and the last from the no-signaling condition in (2.1), so that $I(X; B)_\rho = 0$. Since the inequality holds for all distributions p_X , we can take a supremum to arrive at (4.25). The latter two inequalities in (4.26) and (4.27) follow from well known bounds on conditional mutual information (see, e.g., [Wil16]), and using that systems \hat{A} and X are classical. ■

Definition 16 (Restricted trace distance of assemblages) Let $\{\hat{\rho}_B^{a,x}\}_{a,x}$ and $\{\hat{\theta}_B^{a,x}\}_{a,x}$ be two assemblages. We define the restricted normalized trace distance of assemblages as

$$\Delta^R(\hat{\rho}, \hat{\theta}) := \frac{1}{2} \sup_{p_X} \|\rho_{X\hat{A}B} - \theta_{X\hat{A}B}\|_1, \quad (4.35)$$

where

$$\rho_{X\hat{A}B} := \sum_{x,a} p_X(x)[x] \otimes [a] \otimes \hat{\rho}_B^{a,x}, \quad (4.36)$$

$$\theta_{X\hat{A}B} := \sum_{x,a} p_X(x)[x] \otimes [a] \otimes \hat{\theta}_B^{a,x}. \quad (4.37)$$

Proposition 17 (Metric) The restricted trace distance of assemblages is a metric.

Theorem 18 (Uniform continuity) Let $\{\hat{\rho}_B^{a,x}\}_{a,x}$ and $\{\hat{\theta}_B^{a,x}\}_{a,x}$ be two assemblages such that $\Delta^R(\hat{\rho}, \hat{\theta}) \leq \varepsilon \in [0, 1]$. Then

$$\left| R_S^R(\hat{A}; B)_{\hat{\rho}} - R_S^R(\hat{A}; B)_{\hat{\theta}} \right| \leq \varepsilon \log_2 \min\{|\hat{A}|, |B|\} + g(\varepsilon), \quad (4.38)$$

where $g(\varepsilon) := (\varepsilon + 1) \log_2(\varepsilon + 1) - \varepsilon \log_2 \varepsilon$.

Proposition 19 Let $\varepsilon \in [0, 1]$, and let $\{\hat{\rho}_B^{a,x}\}_{a,x}$ and $\{\hat{\sigma}_B^{a,x}\}_{a,x}$ be assemblages such that $\hat{\sigma} \in \text{LHS}$ and $\Delta^R(\hat{\rho}, \hat{\sigma}) \leq \varepsilon$. Then

$$R_S^R(\hat{A}; B)_{\hat{\rho}} \leq \varepsilon \log_2 \min\{|\hat{A}|, |B|\} + g(\varepsilon). \quad (4.39)$$

Proposition 20 Let $\{\hat{\rho}_B^{a,x}\}_{a,x}$ be an assemblage. Then

$$\sqrt{2 \ln 2} R_S^R(\hat{A}; B)_{\hat{\rho}} \geq \inf_{\{\hat{\sigma}_B^{a,x}\}_{a,x} \in \text{LHS}} \frac{1}{|\mathcal{X}|} \sum_{x,a} \|\hat{\rho}_B^{a,x} - \hat{\sigma}_B^{a,x}\|_1. \quad (4.40)$$

In particular, if $R_S^R(\hat{A}; B)_{\hat{\rho}} = 0$, then $\{\hat{\rho}_B^{a,x}\}_{a,x} \in \text{LHS}$.

5 Conclusion

We provided a definition of relative entropy of steering different from that in [GA15], justifying it based on well grounded information-theoretic and game-theoretic concerns. We showed how this modified relative entropy of steering satisfies several desirable properties, including convexity, steering monotonicity, continuity, and faithfulness. We also considered a restricted relative entropy of steering, which is relevant as a quantifier in an operational setting in which there might be further restrictions on one-way local operations and classical communication, as discussed previously in [KWW16]. The restricted relative entropy of steering is also convex, a steering monotone, continuous, and faithful. Going forward, we suspect that the definitions proposed here will be relevant in applications of steering, such as one-sided device-independent quantum key distribution, but we leave this for future work.

Acknowledgements. EK acknowledges support from the Department of Physics and Astronomy at LSU. MMW acknowledge support from the NSF under Award No. CCF-1350397.

References

- [AC97] Christoph Adami and Nicolas J. Cerf. von Neumann capacity of noisy quantum channels. *Physical Review A*, 56(5):3470–3483, November 1997. arXiv:quant-ph/9609024.
- [BCW⁺12] Cyril Branciard, Eric G. Cavalcanti, Stephen P. Walborn, Valerio Scarani, and Howard M. Wiseman. One-sided device-independent quantum key distribution: Security, feasibility, and the connection with steering. *Physical Review A*, 85(1):010301, January 2012. arXiv:1109.1435.
- [BSST99] Charles H. Bennett, Peter W. Shor, John A. Smolin, and Ashish V. Thapliyal. Entanglement-assisted classical capacity of noisy quantum channels. *Physical Review Letters*, 83(15):3081–3084, October 1999. arXiv:quant-ph/9904023.

- [BSST02] Charles H. Bennett, Peter W. Shor, John A. Smolin, and Ashish V. Thapliyal. Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem. *IEEE Transactions on Information Theory*, 48(10):2637–2655, October 2002. arXiv:quant-ph/0106052.
- [CLL04] Marcos Curty, Maciej Lewenstein, and Norbert Lütkenhaus. Entanglement as a pre-condition for secure quantum key distribution. *Physical Review Letters*, 92(21):217903, May 2004. arXiv:quant-ph/0307151.
- [CS16] Daniel Cavalcanti and Paul Skrzypczyk. Quantum steering: a short review with focus on semidefinite programming. April 2016. arXiv:1604.00501.
- [Dev05] Igor Devetak. The private classical capacity and quantum capacity of a quantum channel. *IEEE Transactions on Information Theory*, 51(1):44–55, January 2005. arXiv:quant-ph/0304127.
- [DHW08] Igor Devetak, Aram W. Harrow, and Andreas Winter. A resource framework for quantum Shannon theory. *IEEE Transactions on Information Theory*, 54(10):4587–4618, October 2008. arXiv:quant-ph/0512015.
- [EPR35] Albert Einstein, Boris Podolsky, and Nathan Rosen. Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 47(10):777–780, May 1935.
- [GA15] Rodrigo Gallego and Leandro Aolita. Resource theory of steering. *Physical Review X*, 5(4):041008, October 2015. arXiv:1409.5804.
- [HHHO05] Karol Horodecki, Michał Horodecki, Paweł Horodecki, and Jonathan Oppenheim. Secure key from bound entanglement. *Physical Review Letters*, 94(16):160502, April 2005. arXiv:quant-ph/0309110.
- [HHHO09] Karol Horodecki, Michał Horodecki, Paweł Horodecki, and Jonathan Oppenheim. General paradigm for distilling classical key from quantum states. *IEEE Transactions on Information Theory*, 55(4):1898–1929, April 2009. arXiv:quant-ph/0506189.
- [Hol98] Alexander S. Holevo. The capacity of the quantum channel with general signal states. *IEEE Transactions on Information Theory*, 44(1):269–273, January 1998. arXiv:quant-ph/9611023.
- [Hol06] Alexander S. Holevo. Multiplicativity of p-norms of completely positive maps and the additivity problem in quantum information theory. *Russian Mathematical Surveys*, 61(2):301–339, 2006.
- [KWW16] Eneet Kaur, Xiaoting Wang, and Mark M. Wilde. Conditional mutual information and quantum steering. December 2016. arXiv:1612.03875.
- [Lin75] Göran Lindblad. Completely positive maps and entropy inequalities. *Communications in Mathematical Physics*, 40(2):147–151, June 1975.
- [Llo97] Seth Lloyd. Capacity of the noisy quantum channel. *Physical Review A*, 55(3):1613–1622, March 1997. arXiv:quant-ph/9604015.

- [OP93] Masanori Ohya and Denes Petz. *Quantum Entropy and Its Use*. Springer, 1993.
- [PLOB15] Stefano Pirandola, Riccardo Laurenza, Carlo Ottaviani, and Leonardo Banchi. October 2015. arXiv:1510.08863v5.
- [Rai01] Eric M. Rains. A semidefinite program for distillable entanglement. *IEEE Transactions on Information Theory*, 47(7):2921–2933, November 2001. arXiv:quant-ph/0008047.
- [Sch35] Erwin Schrödinger. Discussion of probability relations between separated systems. *Mathematical Proceedings of the Cambridge Philosophical Society*, 31(4):555–563, October 1935.
- [Shi15] Maksim Shirokov. Tight continuity bounds for the quantum conditional mutual information, for the Holevo quantity and for capacities of a channel. December 2015. arXiv:1512.09047.
- [Sho02] Peter W. Shor. The quantum channel capacity and coherent information. In *Lecture Notes, MSRI Workshop on Quantum Computation*, 2002.
- [SN96] Benjamin Schumacher and Michael A. Nielsen. Quantum data processing and error correction. *Physical Review A*, 54(4):2629–2635, October 1996. arXiv:quant-ph/9604022.
- [SW97] Benjamin Schumacher and Michael D. Westmoreland. Sending classical information via noisy quantum channels. *Physical Review A*, 56(1):131–138, July 1997.
- [TBR16] Marco Tomamichel, Mario Berta, and Joseph M. Renes. Quantum coding with finite resources. *Nature Communications*, 7:11419, May 2016. arXiv:1504.04617.
- [TGW14] Masahiro Takeoka, Saikat Guha, and Mark M. Wilde. The squashed entanglement of a quantum channel. *IEEE Transactions on Information Theory*, 60(8):4987–4998, August 2014. arXiv:1310.0129.
- [TWW14] Marco Tomamichel, Mark M. Wilde, and Andreas Winter. Strong converse rates for quantum communication. June 2014. arXiv:1406.2946.
- [Uhl77] Armin Uhlmann. Relative entropy and the Wigner-Yanase-Dyson-Lieb concavity in an interpolation theory. *Communications in Mathematical Physics*, 54(1):21–32, 1977.
- [Ume62] Hisaharu Umegaki. Conditional expectations in an operator algebra IV (entropy and information). *Kodai Mathematical Seminar Reports*, 14(2):59–85, 1962.
- [vDGG05] Wim van Dam, Richard D. Gill, and Peter D. Grunwald. The statistical strength of nonlocality proofs. *IEEE Transactions on Information Theory*, 51(8):2812–2835, August 2005. arXiv:quant-ph/0307125.
- [Wil16] Mark M. Wilde. *From Classical to Quantum Shannon Theory*. March 2016. arXiv:1106.1445v7.
- [Win16] Andreas Winter. Tight uniform continuity bounds for quantum entropies: conditional entropy, relative entropy distance and energy constraints. *Communications in Mathematical Physics*, 347(1):291–313, 2016.

- [WJD07] Howard M. Wiseman, S. J. Jones, and Andrew C. Doherty. Steering, entanglement, nonlocality, and the EPR paradox. *Physical Review Letters*, 98(14):140402, April 2007. arXiv:quant-ph/0612147.
- [WTB16] Mark M. Wilde, Marco Tomamichel, and Mario Berta. Converse bounds for private communication over quantum channels. February 2016. arXiv:1602.08898.
- [YHD11] Jon Yard, Patrick Hayden, and Igor Devetak. Quantum broadcast channels. *IEEE Transactions on Information Theory*, 57(10):7147–7162, October 2011. arXiv:quant-ph/0603098.